

MIT-CTP-2341
 hep-th/9408039
 July 1994

Breakdown of the Semi-Classical Approximation at the Black Hole Horizon^{*}

Esko Keski-Vakkuri, Gilad Lifschytz, Samir D. Mathur and Miguel E. Ortiz[†]

*Center for Theoretical Physics, Laboratory for Nuclear Science
 and Department of Physics
 Massachusetts Institute of Technology
 Cambridge MA 02139, USA.
 e-mail: keskivakkuri, gil1, mathur or ortiz @mitlns.mit.edu*

Submitted to: *Physical Review D*

Abstract: The definition of matter states on spacelike hypersurfaces of a 1+1 dimensional black hole spacetime is considered. The effect of small quantum fluctuations of the mass of the black hole due to the quantum nature of the infalling matter is taken into account. It is then shown that the usual approximation of treating the gravitational field as a classical background on which matter is quantized, breaks down near the black hole horizon. Specifically, on any hypersurface that captures both infalling matter near the horizon and Hawking radiation, quantum fluctuations in the background geometry become important, and a semi-classical calculation is inconsistent. An estimate of the size of correlations between the matter and gravity states shows that they are so strong that a fluctuation in the black hole mass of order $e^{-M/M_{Planck}}$ produces a macroscopic change in the matter state.

^{*} This work was supported in part by funds provided by the U.S. Department of Energy (D.O.E.) under cooperative agreement DE-FC02-94ER40818.

[†] Address from September 1994: Dept. of Physics, Tufts University, Medford MA 02155, USA

1 Introduction

Since the original papers of Hawking [1, 2] arguing that black holes should radiate thermally, and that this leads to an apparent loss of information, it has been hoped that investigations of this apparent paradox would lead to a better understanding of quantum gravity. Over the last few years, there has been renewed interest in this general problem. One reason is the construction of 1+1 dimensional models where evaporating black holes can be easily studied [3]. Another reason is the work by 't Hooft [4, 5, 6] suggesting that the black hole evaporation process may not be semiclassical. This idea is based in part on the fact that although Hawking radiation emerges at low frequencies of order M^{-1} at \mathcal{I}^+ , it originates in very high frequency vacuum modes at \mathcal{I}^- and even close to the black hole horizon, the latter frequencies being about e^M times the Planck frequency [7] (here M is the mass of the black hole in Planck units). 't Hooft also argues that if the black hole evaporation process is to be described by unitary evolution, then there should exist large commutators between operators describing infalling matter near the horizon and those describing outgoing Hawking radiation [6] despite the fact that they may be spacelike separated.

Recently¹ Susskind *et. al.* have argued that the information contained in infalling matter could be transferred to the Hawking radiation at the black hole horizon, thus avoiding information loss [9]. A common argument against this possibility is that from the perspective of an infalling observer, who probably sees nothing special at the horizon, there is no mechanism that could account for such a transfer of information. In response, Susskind suggests a breakdown of Lorentz symmetry at large boosts, and a principle of *complementarity* which says that one can make observations either far above the horizon or near the horizon, but somehow it should make no sense to talk of both [9, 10].

The 1+1 dimensional black hole problem including the effects of quantum gravity was recently studied in Ref. [11]. It was found that there are very large commutators between operators at the horizon, and operators at \mathcal{I}^+ measuring the Hawking radiation, agreeing with the earlier work of 't Hooft [6]. Ref. [11] assumes a reflection boundary condition at a strong coupling boundary. Some natural modifications of this boundary condition have been studied recently in [12]. There have been many other studies of quantum gravity on the black hole problem, some of which are listed in [13].

Let us recall the basic structure of the black hole problem [2]. Collapsing matter forms a black hole, which then evaporates by emission of Hawking radiation [1]. The radiation carries away the energy, leaving ‘information’ without energy trapped inside the black hole. The Hawking radiation arises from the production of particle pairs, one member of the pair falling into the horizon and the other member escaping to form the Hawking radiation outside the black hole. The quantum state of the quantum particles outside the black hole is thus not a pure state, and one may compute the entanglement entropy between the particles that fall into the black hole and the particles that escape to infinity. It is possible to carry out such a computation explicitly in the simple 1+1

¹Historically, the greatest champions of this view point have been Page [8] and 't Hooft [4, 5].

dimensional models referred to above. One finds [14, 15] that this entropy equals the quantity expected on the basis of purely thermodynamic arguments [16].

Such calculations are carried out in the semiclassical approximation, where one assumes that the spacetime is a given 1+1 dimensional manifold, and the matter is given by quantum fields propagating on this manifold. How accurate is this description? We wish to examine the viewpoint raised by 't Hooft and Susskind (referred to above) that quantum gravity is important in some sense at the horizon of the black hole. To this end we start with a theory of quantum gravity plus matter, and see how one obtains the semiclassical approximation where gravity is classical but matter is quantum mechanical. The extraction of a semiclassical spacetime from suitable solutions of the Wheeler–DeWitt equation has been studied in [17]. Essentially, one wishes to obtain an approximation where the variables characterizing gravity are ‘fast’ (i.e. the action varies rapidly with change of these variables) and the matter variables are ‘slow’ (i.e. the action varies slowly when they change). This separation hinges on the fact that the gravity action is multiplied by an extra power of the Planck mass squared, compared to the matter variables, and this is a large factor whenever the matter densities are small in comparison to Planck density. We recall that the matter density is indeed low at the horizon of a large black hole (this is just the energy in the Hawking radiation). One might therefore expect the semiclassical approximation to be good at the horizon. It is interesting that this will turn out *not* to be the case, as we shall now show in a 1+1 dimensional model.

It was suggested in [18] that a semiclassical description (*i.e.* where gravity is classical but radiating matter quantum) can break down after sufficient particle production. This suggestion is based on the fact that particle creation creates decoherence [19], but on the other hand an excess of decoherence conflicts with the correlations between position and momentum variables needed for the classical variable [20]. In this paper we investigate this crude proposal and find that there is indeed a sense in which the semiclassical approximation breaks down near a black hole. It turns out that the presence of the horizon is crucial to this phenomenon, so what we observe here is really a property of black holes.

Since in black hole physics one is interested in concepts like entropy, information, and unitarity of states, it is appropriate to use a language where one deals with ‘states’ or ‘wavefunctionals’ on spacelike hypersurfaces, instead of considering functional integrals or correlation functions over a coordinate region of spacetime. In this description, the dynamical degrees of freedom are 1-geometries, and it is more fundamental to speak of the state of matter on a 1-geometry than on an entire spacetime. Thus, we will need to study the canonical formulation of 1+1 dimensional dilaton gravity. Recall that in this theory the gravity sector contains both the metric and an additional scalar field, the dilaton, which together define a 1-geometry. The space of all possible 1-geometries is called superspace. We assume that our theory of quantum gravity plus matter is described by some form of Wheeler–DeWitt equation [21], which enforces the Hamiltonian constraint on wavefunctionals in superspace. For dilaton gravity alone, a point of superspace is given by the fields $\{\rho(x), \phi(x)\}$. Here we have assumed the notation

that the metric along the 1-dimensional geometry is $ds^2 = e^{2\rho} dx^2$, and ϕ is the dilaton. One of the constraints on the wavefunctionals is the diffeomorphism invariance in the coordinate x . Using this invariance we may reduce the description of superspace so that different points just consist of intrinsically different 1-geometries. More precisely, choose any value of ϕ , say ϕ_0 . Let s denote the proper distance along the 1-geometry measured from the point where $\phi = \phi_0$, with s positive in the direction where ϕ decreases. The function $\phi(s)$ along the 1-geometry describes the intrinsic structure of the 1-geometry, and is invariant under spatial diffeomorphisms (we have assumed here for simplicity that ϕ is a monotonic function along the 1-geometry, and that the value ϕ_0 appears at some point along the 1-geometry). Loosely speaking, we may regard superspace as the space of all such functions $\phi(s)$ (for a spacetime with boundary, this description must be supplemented with an embedding condition at the boundary).

Let us now consider the presence of a massless scalar field $f(x)$. Points of superspace now are described by $\{\phi(s), f(\phi(s))\}$, and wavefunctionals on this space, $\Psi[\phi(s), f(\phi(s))]$, satisfy the Wheeler–DeWitt equation

$$(H_{\text{gravity}} + H_{\text{matter}})\Psi[\phi(s), f(\phi)] = 0. \quad (1)$$

We are now faced with the question: How do we obtain the semiclassical limit of quantum gravity, starting from some theory of quantum gravity plus matter? At the present point we have only 1-geometries in the description, and we have to examine how the 1+1 dimensional spacetime emerges in some approximation from $\Psi[\phi(s), f(\phi)]$. Obtaining a 1+1 dimensional spacetime has been called the ‘problem of time’ in quantum gravity, and considerable work has been done on the semiclassical approximation of gravity as a solution to this problem [17]. We wish to reopen this discussion in the context of black hole physics.

In mathematical terms, we have $\Psi[\phi(s), f(\phi(s))]$ giving the complete description of matter plus gravity. What is the state of matter on a time-slice? If we are given a classical 1+1 spacetime, then a time-slice is given by an intrinsic 1-geometry $\phi(s)$ (plus a boundary condition at infinity). Thus the matter wavefunctional on a time-slice $\phi(s)$ should be given by

$$\Psi_{\phi(s)}[f(\phi(s))] \equiv \Psi[\phi(s), f(\phi(s))] , \quad (2)$$

The semiclassical approximation then consists of approximating the full solution of the Wheeler–DeWitt equation by the product of a semiclassical functional of the gravitational variables alone, times a matter part which is taken to be a solution

$$\psi_{\phi(s)}^{\mathcal{M}}[f(\phi(s))] \quad (3)$$

of the functional Schrödinger equation on some mean spacetime \mathcal{M} (here the function $\phi(s)$ is like a generalized time coordinate on \mathcal{M}). If any quantum field theory on curved spacetime calculation using (3) can be used to approximate the result obtained using the exact solution of the Wheeler–DeWitt equation of (2), then we say that the semiclassical approximation is good. On the other hand, if this approximation fails to

work, we conclude that quantum fluctuations in geometry are important to whichever question it is that we wished to answer.

For the black hole problem, it is appropriate to make a separation between the matter regarded as forming the black hole, denoted by $F(\phi(s))$, and all other matter $f(\phi(s))$. It is then more natural to regard $F(\phi(s))$ as part of the gravitational degrees of freedom, and it is certainly regarded as a classical background field in the derivation of Hawking radiation using the semiclassical approximation. In this situation we must be more precise about what we require for the semiclassical approximation to be good. Assume that the black hole is formed by the collapse of some wavepacket of matter F , into a region smaller than the Schwarzschild radius. We note that the energy of this matter wavepacket cannot be exactly M , because an eigenstate of energy would not evolve at all over time in the manner needed to describe the collapsing packet. In fact, since the matter will be localized to within the Schwarzschild radius M , there will be a momentum uncertainty much greater than $1/M$ in Planck units, which leads to an energy uncertainty which must also be much larger than $1/M$. This uncertainty is still quite small, but should nevertheless not be ignored. The different possible energy values in this range $(M, M + \Delta M)$ where $\Delta M \gg 1/M$, will give different semiclassical spacetimes. For the semiclassical approximation to be good for any given computation, it must be independent of which of the slightly different spacetimes is chosen. Conversely, if the difference in any quantity of interest becomes significant when evaluated on different spacetimes in the above mass range, then we cannot use a mean 2-geometry to describe physics, and we should say that the semiclassical approximation is not good²

Casting this problem in the language of the preceeding paragraphs, we must ask whether the wavefunctional of matter from the full quantum solution of the Wheeler–DeWitt equation is well approximated by working on a fixed spacetime \mathcal{M} of mass M and ignoring the uncertainty ΔM in M . Now, suppose that the semiclassical approximation were a good one when describing the state of matter on a given time-slice $\phi(s)$. If we consider the different matter states that are obtained on $\phi(s)$ by taking different values for M , which cannot be clearly distinguished because we are averaging over the fluctuations in geometry, then these states should not be ‘too different’ if there is to be an unambiguous definition of the state on the time-slice. This is a minimal requirement for a semiclassical calculation to be a good approximation to $\Psi[\phi(s), F(\phi(s)), f(\phi(s))]$.

Let the state of quantized matter obtained by working on \mathcal{M} be $\psi_{\phi(s)}^{\mathcal{M}}[f(\phi(s))]$, where in \mathcal{M} the energy of the infalling matter is M . This is a state in the Schrödinger representation, and thus depends on the time-slice specified by the function $\phi(s)$ (plus boundary condition). At slices corresponding to early times (i.e. near \mathcal{I}^- , before the black hole formed) for all spacetimes with mass M in the range $(M, M + \Delta M)$, we fix the matter state to be approximately the same in each spacetime. In terms of a natural inner product relating states on a common 1-geometry in different spacetimes

²The role of fluctuations in the mass of the infalling matter was also discussed in [11]. Generally, fluctuations in geometry can also arise from other sources, but we shall ignore these here.

(which we define in this paper), this means that

$$\langle \psi_{\phi(s)}^{\mathcal{M}} | \psi_{\phi(s)}^{\bar{\mathcal{M}}} \rangle \approx 1 \quad (4)$$

on these early time slices, where $\bar{\mathcal{M}}$ is a spacetime with mass \bar{M} in the above range. On each spacetime the matter state evolves in the Schrödinger picture in different ways, so that the inner product (4) will not be the same on all slices. For the semiclassical approximation to be good at any given slice, we need that (4) hold on that slice.

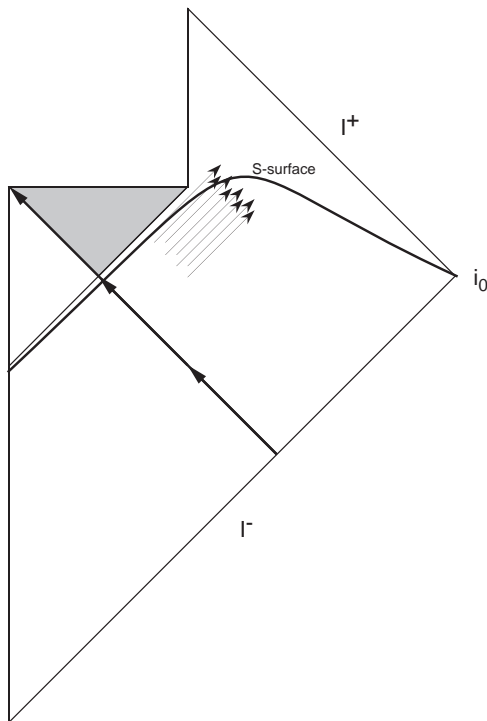


Figure 1: An example of an S-surface, shown in an evaporating black hole spacetime.

Having fixed the matter states on different spacetimes so that they are very similar at early times, we analyse later time slices to check that this property still holds. Any slice is taken to start at some fixed base point near spatial infinity. Consider now a slice that moves up in time near \mathcal{I}^+ to capture some fraction of the Hawking radiation. The slice then comes to the vicinity of the horizon, and then moves close to the horizon, so as to reach early advanced times before entering the strong coupling domain (see Fig. 1). The importance of such slices to the black hole paradox has been emphasized by Preskill [24] and Susskind *et. al.* [9] in their arguments relating to information bleaching and to the principle of black hole complementarity. Susskind *et. al.* conjectured that the large Lorentz boost between the two portions of the slice should lead to a problem in the semiclassical description of a black hole. Slices of this type have also been used in

the literature as part of a complete spacelike slicing of spacetime, that stays outside the horizon of the black hole [22] and captures the Hawking radiation, and on which semiclassical physics should therefore apply. For these surfaces, which we shall refer to as S-surfaces, we shall show in this paper that it is no longer the case that matter states are approximately the same for different background spacetimes. Indeed, even for $|M - \bar{M}| \sim e^{-M}$ we find that on a 1-geometry $\phi(s)$ of this type,

$$\langle \psi_{\phi(s)}^{\mathcal{M}} | \psi_{\phi(s)}^{\bar{\mathcal{M}}} \rangle \approx 0.$$

As was argued above, the fluctuations in the mass of the hole must be at least of order $\Delta M > 1/M$, so we see that the state of matter on such slices is very ill defined because of the fluctuations in geometry. This shows that at least one natural quantity that we wish to consider in black hole physics, the state of matter on what we have termed an S-surface, is not given well by the semiclassical approximation.

The plan of this paper is the following. In section 2 we review the CGHS model, and give some relevant scales. In section 3 we study the embedding of 1-geometries in different 1+1 dimensional semiclassical spacetimes. In section 4 we compare states of matter on the same 1-geometry, but in different spacetimes. Section 5 is a general discussion of the meaning of these results and of possible connections to other work.

2 A review of the CGHS model

There follows a quick review of the CGHS model [3], with reference to the RST model [23] which includes back-reaction and defines some relevant scales in the CGHS solution. Although all calculations in this paper are for a CGHS black hole, the general features of the results that are derived are expected to apply equally well to other black hole models in two and four dimensions.

The Lagrangian for two dimensional string-inspired dilaton gravity is

$$S_G = \frac{1}{2\pi} \int dx dt \sqrt{-g} e^{-2\phi} \left[R + 4(\nabla\phi)^2 + 4\lambda^2 \right] \quad (5)$$

where $\phi(x)$ is the dilaton field and λ is a parameter analogous to the Planck scale. Writing

$$ds^2 = -e^{2\rho} dx^+ dx^-$$

where $x^\pm = t \pm x$ are referred to as Kruskal coordinates, (5) has static black hole solutions

$$e^{-2\rho} = e^{-2\phi} = \frac{M}{\lambda} - \lambda^2 x^+ x^- \quad (6)$$

and a linear dilaton vacuum (LDV) solution with $M = 0$. More interesting is the solution obtained when (5) is coupled to conformal matter,

$$S = S_G - \frac{1}{4\pi} \int dx dt \sqrt{-g} (\nabla f)^2,$$

where f is a massless scalar field. A left moving shock wave in f giving rise to a stress tensor

$$\frac{1}{2}\partial_+ f \partial_+ f = M\delta(x^+ - 1/\lambda)$$

yields a solution

$$e^{-2\rho} = e^{-2\phi} = -\frac{M}{\lambda}(\lambda x^+ - 1)\Theta(x^+ - 1/\lambda) - \lambda^2 x^+ x^- \quad (7)$$

representing the formation of a black hole of mass M/λ in Planck units (the Penrose diagram for this solution is shown in Fig. 2). For $\lambda x^+ < 1$ (region I), the solution is simply the LDV, whereas the solution for $\lambda x^+ > 1$ (region II),

$$e^{-2\rho} = e^{-2\phi} = \frac{M}{\lambda} - \lambda x^+ \left(\lambda x^- + \frac{M}{\lambda} \right)$$

is a black hole with an event horizon at $\lambda x^- = -M/\lambda$.

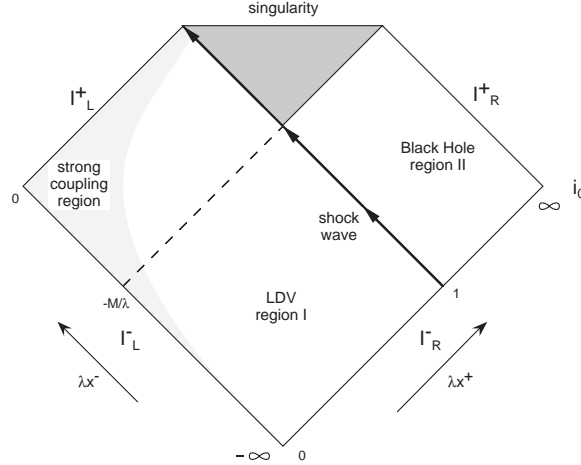


Figure 2: The Penrose diagram of the CGHS solution.

It is possible to define asymptotically flat coordinates in both regions I and II. In region I, we define

$$\lambda x^+ = e^{\lambda y^+}, \quad \lambda x^- = -\frac{M}{\lambda} e^{-\lambda y^-} \quad (8)$$

and in region II we introduce the “tortoise” coordinates $\lambda\sigma^\pm$:

$$\lambda x^+ = e^{\lambda\sigma^+}, \quad \lambda x^- + \frac{M}{\lambda} = -e^{-\lambda\sigma^-} . \quad (9)$$

The coordinate y^- is used to define right moving modes at \mathcal{I}_L^- . To define left moving modes at \mathcal{I}_R^- we can use either y^+ or σ^+ . As (8) and (9) tell us, both coordinates can

be extended to $I \cup II$ so that $y^+ = \sigma^+$. It is easy to see that as $\sigma \rightarrow \infty$ or as $y \rightarrow \infty$, $\rho \rightarrow -\infty$. Notice also that e^ϕ plays the role of the gravitational coupling constant in this theory. It is generally believed that semiclassical theory is reliable in regions where this quantity is small. At infinity $e^\phi \rightarrow 0$, and so this is a region of very weak coupling. Even at the horizon, $e^\phi = \sqrt{\lambda/M}$ is small provided that the mass of the black hole is large in Planck units ($M/\lambda \gg 1$). This is assumed to be the case in all calculations so that the weak coupling region extends well inside the black hole horizon.

One virtue of this two dimensional model is that it is straightforward to include the effects of backreaction by adding counterterms to the action S . This was first done by CGHS, but a more tractable model was introduced by RST who found an analytic solution for the metric of an evaporating black hole. However, the RST model still exhibits all the usual paradoxes associated with black hole evaporation (for a review see [24, 25]).

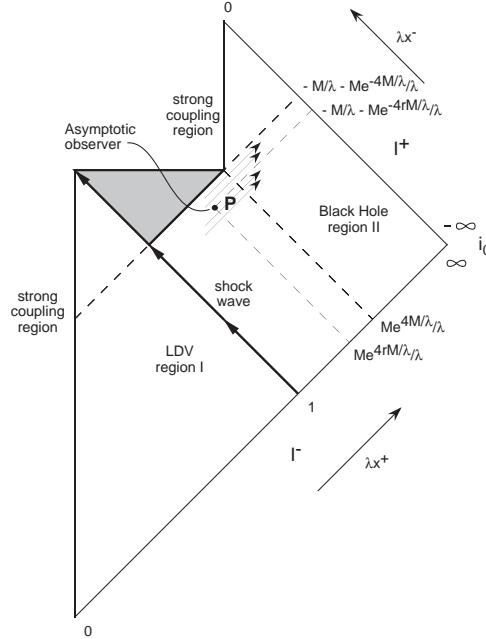


Figure 3: The Penrose diagram of the RST solution with some approximate scales shown.

Although we will carry out our calculations in the simpler CGHS model, the RST solution (whose Penrose diagram is shown in Fig. 3), is a useful guide for identifying certain scales in the evaporation process. These can be usefully carried over to a study of the CGHS solution, and serve to determine the portion of that solution that is unaffected by backreaction: The time scale of evaporation of the hole as measured by an asymptotic observer is $t_E \sim 4M$ in Planck units; the value of x^- at which a proportion r of the total Hawking radiation reaches \mathcal{I}^+ is $\lambda x_P^- = -M(1 + e^{-4rM/\lambda})/\lambda$

(by this we mean that the Hawking radiation to the right of this value carries energy rM); the value of x^+ , for $x^- = x_P^-$, which corresponds to a point well outside the hole, in the sense that the curvature is weak and the components of the stress tensor are small is $\lambda x_P^+ = M e^{4rM/\lambda}/\lambda$ provided that $\lambda x^+ > e^{2M/\lambda}$. On the basis of these scales, we can define a point P at $(\lambda x_P^+, \lambda x_P^-)$ as defined above, located just outside the black hole, in the asymptotically flat region, and to the left of a proportion r of the Hawking radiation.

3 Embedding of 1-geometries

In this section, we shall compare how a certain spacelike hypersurface Σ may be embedded in collapsing black hole spacetimes (7) of masses M (denoted by \mathcal{M}) and $\bar{M} = M + \Delta M$ (denoted by $\bar{\mathcal{M}}$), where ΔM is a fluctuation of at most Planck size.

In 1+1 dimensional dilaton gravity models an invariant definition of a 1-geometry is provided by the value of the dilaton field $\phi(s)$ as a function of the proper distance s along the 1-geometry, measured from some fixed reference point. For spacetimes with boundary, such as the black hole geometries in the CGHS model, this reference point may be replaced by information about how the 1-geometry is embedded at infinity. It is natural to regard asymptotic infinity as a region where hypersurfaces can be nailed down by external observers who are not a part of the quantum system we are considering. We impose the condition that 1-geometries in different spacetimes should be indistinguishable for these asymptotic observers, ensuring that the semiclassical approximation holds for these observers. This condition and the function $\phi(s)$ are enough to define a unique map of Σ from \mathcal{M} to $\bar{\mathcal{M}}$.

It is important to point out at this stage that it is possible that this map is not well defined for some $\bar{\mathcal{M}}$, in the sense that there may exist no spacelike hypersurface in $\bar{\mathcal{M}}$ with the required properties. For the surfaces we consider, this issue does not arise. Further, it can be argued that there is no important effect of this phenomenon on the state of the matter fields, at least as long as one is away from strong curvature regions. (To see this it is helpful to use the explicit quantum gravity wavefunction for dilaton gravity given in [26]). For this reason we shall ignore all spacetimes $\bar{\mathcal{M}}$ where Σ does not fit.

Given an equation for $\Sigma_{\mathcal{M}}$,

$$\lambda x^- = f(\lambda x^+)$$

and expressions for $\rho(x^+, x^-)$ and $\phi(x^+, x^-)$ in \mathcal{M} and $\bar{\rho}(\bar{x}^+, \bar{x}^-)$ and $\bar{\phi}(\bar{x}^+, \bar{x}^-)$ in $\bar{\mathcal{M}}$, we determine the corresponding equation for $\Sigma_{\bar{\mathcal{M}}}$,

$$\lambda \bar{x}^- = \bar{f}(\lambda \bar{x}^+)$$

by requiring that $\phi(s) = \bar{\phi}(\bar{s})$ and similarly $d\phi/ds(s) = d\bar{\phi}/d\bar{s}(\bar{s})$ (it is if these equations have no real solution for a given $\bar{\mathcal{M}}$ that we say that Σ does not fit in $\bar{\mathcal{M}}$). These conditions require one boundary condition which fixes $\Sigma_{\bar{\mathcal{M}}}$ at infinity, and this may be

chosen in such a way that the equations for $\Sigma_{\mathcal{M}}$ and $\Sigma_{\bar{\mathcal{M}}}$ are the same in asymptotically flat (tortoise) coordinates sufficiently far from the black hole.

We shall demonstrate that while most surfaces embed in very slightly different ways in spacetimes \mathcal{M} and $\bar{\mathcal{M}}$ with masses differing only at the Planck scale, there is a special class of surfaces for which this is not true (what we mean by embeddings being different will be discussed later). These are the S-surfaces which catch both the Hawking radiation (the Hawking pairs reaching \mathcal{I}^+ , but not those ending up at the singularity) and the in-falling matter near the horizon (see Fig. 1). It is useful to give an example of such surfaces. A straight line in Kruskal coordinates x^\pm going through a point $P \sim (Me^{4rM/\lambda}/\lambda, -M(1 + e^{-4rM/\lambda}))$, is a line of this type, catching a proportion r of the outgoing Hawking radiation, provided the slope of the line is extremely small – of order $e^{-8rM/\lambda}$. The smallness of this parameter will play an important role in our discussion. Although the line is straight in Kruskal coordinates, it will, of course, look bent in the Penrose diagram, ending up at i_0 . Far from the horizon, these lines are lines of constant Schwarzschild time $\lambda t = 4rM$, giving an interpretation for minus one half the logarithm of the slope in terms of the time at infinity.

It is worth pointing out that the map from a surface $\Sigma_{\mathcal{M}}$ in \mathcal{M} to the corresponding surface $\Sigma_{\bar{\mathcal{M}}}$ in $\bar{\mathcal{M}}$ defines a map from any point Q on $\Sigma_{\mathcal{M}}$ to a point \bar{Q}_Σ on $\Sigma_{\bar{\mathcal{M}}}$ in $\bar{\mathcal{M}}$. Any other choice of surface $\Xi_{\mathcal{M}}$ in \mathcal{M} passing through Q maps Q to a different point \bar{Q}_Ξ in $\bar{\mathcal{M}}$. This uncertainty in the location of a point \bar{Q} in $\bar{\mathcal{M}}$ gives a geometric way of defining the fluctuations in geometry around Q . Generally, we may expect all the images of Q in $\bar{\mathcal{M}}$ to lie within a small region of Planck size. However, we shall see below that this is not the case near a black hole horizon.

3.1 Basic Equations

Here we present the basic equations describing the embedding of Σ . In a collapsing black hole manifold \mathcal{M} of mass M (7), it is convenient to define Σ as

$$\lambda x^- = f(\lambda x^+) - M/\lambda.$$

If we use Kruskal coordinates \bar{x}^\pm to describe Σ in a black hole manifold $\bar{\mathcal{M}}$ of mass \bar{M} as

$$\lambda \bar{x}^- = \bar{f}(\lambda \bar{x}^+) - \bar{M}/\lambda,$$

then in region II of (7)

$$\frac{M}{\lambda} - \lambda x^+ f(\lambda x^+) = \frac{\bar{M}}{\lambda} - \lambda \bar{x}^+ \bar{f}(\lambda \bar{x}^+) \quad (10)$$

$$\frac{f + \lambda x^+ f'}{\sqrt{-f'}} = \frac{\bar{f} + \lambda \bar{x}^+ \bar{f}'}{\sqrt{-\bar{f}'}} \quad (11)$$

where prime denotes a derivative with respect to the argument. The first equation is the requirement of equal $\phi(s)$ and the second of equal $d\phi/ds(s)$.

Once we identify the embedding of Σ in $\bar{\mathcal{M}}$, we can then identify points in both spacetimes by the value of s on Σ . This identification may be described by the function $\bar{x}^+(x^+)$ between coordinates on Σ in each of the spacetimes. To solve the equations (10) and (11), for $\bar{x}^+(x^+)$, differentiate (10) by x^+ and divide by (11), to get

$$\frac{d\bar{x}^+}{dx^+} = \frac{\sqrt{-f'}}{\sqrt{-\bar{f}'}}.$$

Another combination of these equations gives

$$\sqrt{-\bar{f}'} = \frac{-(f + \lambda x^+ f') \pm \sqrt{(f - \lambda x^+ f')^2 - 4\Delta M f'/\lambda}}{2\lambda \bar{x}^+ \sqrt{-f'}}$$

where $\Delta M = \bar{M} - M$. Combining both equations,

$$\ln(\lambda \bar{x}^+) = 2 \int d(\lambda x^+) \frac{f'}{(f + \lambda x^+ f') \mp \sqrt{(f - \lambda x^+ f')^2 - 4\Delta M f'/\lambda}} \quad (12)$$

which is a general expression for $\bar{x}^+(x^+)$ for any Σ . Similarly, if we label the one geometry by $\lambda x^+ = g(z^-)$ where $z^- = \lambda x^- + M/\lambda$ (using the notation $g = f^{-1}$), we find an analogous expression for $\bar{x}^-(x^-)$:

$$\ln(\lambda \bar{x}^- + \bar{M}/\lambda) = 2 \int dz^- \frac{g'}{(g + z^- g') \mp \sqrt{(g - z^- g')^2 - 4\Delta M g'/\lambda}} \quad , \quad (13)$$

In (12) and (13), the sign of the square root is determined by requiring that as ΔM tends to zero we get $\bar{x}^\pm = x^\pm$. From these equations one can construct the corresponding one geometry in $\bar{\mathcal{M}}$. In order for the solution to make sense, the expressions inside the square root must be positive. This condition is a manifestation of the fitting problem mentioned above.

3.2 A large shift for straight lines

For simplicity, we focus our attention on lines that are straight in the Kruskal coordinates x^\pm . Below we present a quick analysis of the embedding of these 1-geometries in neighbouring spacetimes. In the next subsection a more detailed treatment will be given.

Consider the line Σ defined in \mathcal{M} by the equation

$$\lambda x^- = f(\lambda x^+) - \frac{M}{\lambda} = -\alpha^2 \lambda x^+ + b.$$

It is easy to see that as a consequence of (10) and (11), the function $\bar{f}(\lambda \bar{x}^+)$ describing the deformed line in Kruskal coordinates on $\bar{\mathcal{M}}$ must also be linear. This is a helpful simplification. Let us write the equation for Σ in $\bar{\mathcal{M}}$ as

$$\lambda \bar{x}^- = \bar{f}(\lambda \bar{x}^+) - \frac{\bar{M}}{\lambda} = -\bar{\alpha}^2 \lambda \bar{x}^+ + \bar{b}$$

The parameters b and \bar{b} are related by

$$\frac{(\bar{b} + \bar{M}/\lambda)^2}{\bar{\alpha}^2} = \frac{4\Delta M}{\lambda} + \frac{(b + M/\lambda)^2}{\alpha^2} \quad (14)$$

It is useful to define another quantity δ , so that Σ crosses the shock wave, ($\lambda x^+ = 1$) in \mathcal{M} at $\lambda x^- = -M/\lambda - \delta$ (*i.e.* $\delta = \alpha^2 - b - M/\lambda$). We then find from equation (12) that

$$2\bar{\alpha}\lambda\bar{x}^+ = 2\alpha\lambda x^+ + \delta/\alpha - \alpha \pm \sqrt{(\alpha^2 - \delta)^2/\alpha^2 + 4\Delta M/\lambda}$$

We still have a free parameter $\bar{\alpha}$. The way to fix it is by imposing the condition that Σ should be the same for an asymptotic observer at infinity, meaning that as expressed in tortoise coordinates σ or $\bar{\sigma}$, Σ should have the same functional form up to unobservable (Planck scale) perturbations. This may be achieved, as we will see later, simply by picking a point on Σ in \mathcal{M} , call it x_0^+ , and demanding that both lines have the same value of ϕ at the point $x^+ = \bar{x}^+ = x_0^+$. Then

$$\bar{\alpha} = \alpha + \frac{\delta/\alpha - \alpha \pm \sqrt{(\alpha^2 - \delta)^2/\alpha^2 + 4\Delta M/\lambda}}{2x_0^+} \quad (15)$$

Taking $x_0 \rightarrow \infty$ fixes the line at infinity. The result does not depend on whether we take $x_0 \rightarrow \infty$ or just take it to be in the asymptotic region $x_0 > Me^{2M/\lambda}/\lambda$.

We can actually derive some quite general conclusions about how the embedding of Σ changes from \mathcal{M} to $\bar{\mathcal{M}}$ from (14) and (15). Let us split the possible Σ 's into three simple cases, for any value of α and δ (recall that $|\Delta M/\lambda| < 1$):

1. $(\alpha^2 - \delta)^2/\alpha^2 \gg 4\Delta M/\lambda$

In this case

$$\bar{\alpha} = \alpha$$

and

$$\lambda\bar{x}^+ = \lambda x^+ + \frac{\Delta M}{\lambda(\alpha^2 - \delta)} \quad .$$

2. $(\alpha^2 - \delta)^2/\alpha^2 \ll 4\Delta M/\lambda$

For $\Delta M/\lambda \geq 0$ (this is taken to avoid fitting problems)

$$\bar{\alpha} = \alpha$$

and

$$\lambda\bar{x}^+ = \lambda x^+ \pm \frac{\sqrt{\Delta M/\lambda}}{\alpha} \quad .$$

3. $(\alpha^2 - \delta)^2/\alpha^2 \sim 4\Delta M/\lambda$

Again $\Delta M/\lambda \geq 0$, and we find a similar result

$$\alpha = \bar{\alpha}$$

and

$$\lambda \bar{x}^+ \sim \lambda x^+ \pm \frac{\sqrt{\Delta M/\lambda}}{\alpha}.$$

In the last two cases the sign \pm depends on the sign of $\alpha^2 - \delta$.

The above results all show that the slope $\bar{\alpha}$ of the line in $\bar{\mathcal{M}}$ is virtually identical to the slope α in \mathcal{M} (identical in the limit $x_0 \rightarrow \infty$). It is also the case that the position of the line in the x^- direction is almost the same in \mathcal{M} and $\bar{\mathcal{M}}$. However, for lines with small values of α and δ , there is a large shift in the location of the line in the x^+ direction in $\bar{\mathcal{M}}$ relative to its position in \mathcal{M} . The lines for which this effect occurs are precisely the S-surfaces that we have discussed above. These were defined to have $\alpha^2 \sim M e^{-8rM/\lambda}/\lambda$, and $0 \leq \delta \leq M e^{-4rM/\lambda}/\lambda$, which are both small enough to compensate for the ΔM in the numerator in the expressions above. The large shift, and the fact that it occurs only for a very specific class of lines, precisely the S-surfaces which capture both a reasonable proportion of the Hawking radiation and the infalling matter (see Fig. 1), is the fundamental result behind the arguments presented in this paper. The fact that only a special class of lines exhibit this effect is reassuring, as it means that any effects that are a consequence of this shift can only be present close to the black hole horizon.

3.3 Complete hypersurfaces

So far we have not taken the hypersurfaces to be complete, *i.e.*, we have not done the full calculation of continuing them to the LDV and finishing at infinity in the strong coupling regime. We will now perform the full calculation for a certain class of hypersurfaces. They will provide us a convenient example (for calculational purposes) for use in section 4, where we will discuss the implications of the large shift on the time evolution of matter states.

We choose, for convenience, to work with a class of hypersurfaces that all have $d\phi/ds = -\lambda$:

$$\lambda x^- = \begin{cases} -\alpha^2 \lambda x^+ - 2\alpha \sqrt{\frac{M}{\lambda}} - \frac{M}{\lambda} & (\lambda x^+ \geq 1) \\ -\left(\alpha + \sqrt{\frac{M}{\lambda}}\right)^2 \lambda x^+ & (\lambda x^+ \leq 1) \end{cases}. \quad (16)$$

These lines are of type 1 ($(\alpha^2 - \delta)^2/\alpha^2 \gg 4\Delta M/\lambda$) discussed in section 3.2. They have one free parameter, the slope α^2 . At spacelike infinity, these lines are approximately constant Schwarzschild time lines, $\sigma^0 = -\ln \alpha$, and for different values of α , they provide a foliation of spacetime in a way often discussed in the literature [22] in the context of the black hole puzzle. They always stay outside the event horizon, and they cross the shock wave at a Kruskal distance $\delta = 2\alpha\sqrt{M/\lambda} + \alpha^2$ from the horizon. After crossing the shock wave they continue to the strong coupling region. For an early time Cauchy surface, the parameter α^2 is arbitrarily large ($\alpha^2 \rightarrow \infty$ would make the lines

approach \mathcal{I}^-). As α^2 becomes smaller, the lines move closer to the event horizon. Finally, as $\alpha^2 \rightarrow 0$, the upper segment asymptotes to \mathcal{I}^+ and to the segment of the event horizon above the shock wave. This is illustrated in Fig. 4. We are mostly interested in the S-surfaces that catch a ratio r of the Hawking radiation emitted by the black hole, which fixes the value of α . For r not too close to 1, the S-surfaces are well within the weak coupling region.

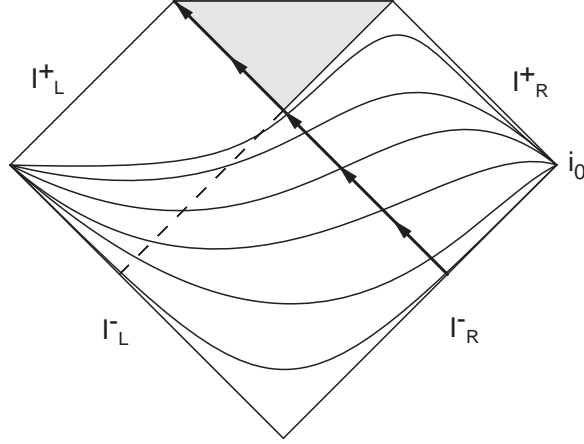


Figure 4: Examples of the complete slices of Sec. 3.3.

We want to find the location of the above lines in a black hole background with a mass $\bar{M} = M + \Delta M$. It is easy to see that in the new background, the lines

$$\lambda \bar{x}^- = \begin{cases} -\bar{\alpha}^2 \lambda \bar{x}^+ - 2\bar{\alpha} \sqrt{\frac{\bar{M}}{\lambda}} - \frac{\bar{M}}{\lambda} & (\lambda \bar{x}^+ \geq 1) \\ -\left(\bar{\alpha} + \sqrt{\frac{\bar{M}}{\lambda}}\right)^2 \lambda \bar{x}^+ & (\lambda \bar{x}^+ \leq 1) \end{cases}. \quad (17)$$

also satisfy $d\phi/ds \equiv -\lambda$. We only need to identify the new slope $\bar{\alpha}^2$ in terms of the old one, and as before, this is given by the boundary conditions at infinity. Requiring $\lambda \bar{\sigma}_0^+ = \lambda \sigma_0^+$, where σ^+ is the tortoise coordinate defined in (9), yields

$$\bar{\alpha} = \alpha + \left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\bar{M}}{\lambda}} \right) e^{-\lambda \sigma_0^+}.$$

If we also want to require $\lambda \bar{\sigma}_0^- = \lambda \sigma_0^-$, we need to do the fixing at infinity, which of course sets

$$\bar{\alpha} = \alpha$$

After fixing the σ^\pm coordinates at infinity, we may check that $\bar{\sigma}^\pm$ and σ^\pm do not differ appreciably as we approach the point P (still considered to be in the asymptotic region) along an S-surface. Taking $\alpha \sim e^{-4rM/\lambda}$ and P to be at $\lambda x_P^+ \sim Me^{4rM/\lambda}/\lambda$, $\lambda x_P^- \sim -M/\lambda - Me^{-4rM/\lambda}/\lambda$ as before, we find that at P

$$\begin{aligned}\lambda \bar{\sigma}_P^+ - \lambda \sigma_P^+ &\approx -\frac{\Delta M}{2M} \left(\frac{\lambda}{M} \right)^{1/2} \\ \lambda \bar{\sigma}_P^- - \lambda \sigma_P^- &\approx -\frac{\Delta M}{2M} \ ,\end{aligned}\tag{18}$$

which is a small deviation. We conclude that if we had fixed the surface at P instead of infinity, all results would be qualitatively unchanged, as one would expect.

We can now compute the relationship between λx^+ and $\lambda \bar{x}^+$. As we saw in the previous subsection, the points in the original line get “shifted” by a large amount in the new line. It is easy to see that

$$\lambda \bar{x}^+ = \lambda x^+ + \frac{1}{\alpha} \left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{M}{\lambda}} \right)\tag{19}$$

$$\approx \lambda x^+ - \frac{\Delta M}{2\lambda\alpha} \sqrt{\frac{\lambda}{M}} \ .\tag{20}$$

For instance, for $\alpha \sim e^{-4rM/\lambda}$, $\Delta M/\lambda \sim \lambda/M$ the shift is of the order of

$$\lambda \bar{x}^+ - \lambda x^+ \sim -\frac{1}{2} \left(\frac{\lambda}{M} \right)^{3/2} e^{4rM/\lambda} \ ,$$

which is huge. Even for $\Delta M/\lambda \sim e^{-M/\lambda}$, the shift can be extremely large. As we will see in section 4, instead of the relations $\lambda \bar{x}^+ = \lambda \bar{x}^+(\lambda x^+)$, we will be interested in the induced relations between the asymptotically flat coordinates $\lambda \bar{\sigma}^+$ and $\lambda \sigma^+$, and $\lambda \bar{y}^-$ and λy^- . A huge shift in the Kruskal coordinate close to the shock wave will correspond to a big shift in the coordinate $\lambda \sigma^+$, in which the metric is flat at \mathcal{I}_R^- . As a consequence, the relation between $\lambda \bar{\sigma}^+$ and $\lambda \sigma^+$ is nonlinear, as we will discuss in the next section.

Finally, let us mention an immediate consequence of this large shift in the x^+ direction. The map of an S-surface from \mathcal{M} to $\bar{\mathcal{M}}$ induces a map from a point Q close to the horizon to a point \bar{Q} which is shifted a long way up the horizon in terms of Kruskal coordinates. A similar map induced by other surfaces through Q which are not S-surfaces will not shift \bar{Q} by a large amount. We therefore see the presence of large quantum fluctuations near the horizon in the position of \bar{Q} in the sense defined above. These large fluctuations are already a somewhat unexpected result.

4 The state of matter on Σ

We have seen in the previous section in some detail the large shift that occurs in the x^+ direction when we map a S-surface Σ from a black hole spacetime \mathcal{M} to one with a mass

which differs from \mathcal{M} by an extremely small amount, even compared with the Planck scale. This appears to be a large effect, capable of seriously impairing the definition of a unique quantum matter state on Σ in a semiclassical way. There are, however, many large scales in the black hole problem, and it is premature to draw conclusions from the appearance of this large shift in the Kruskal coordinates, without verifying that there is a corresponding shift in physical (coordinate invariant) quantities. An absolute measure of the shift is given by the asymptotic tortoise coordinate σ^+ at \mathcal{I}_R^- . The exponential relationship between x^+ and σ^+ implies that the shift is of Planck size for an x^+ far from the shock wave ($x^+/x_P^+ \sim 1$, where x_P is again as defined at the end of section 2), and there is no reason to expect this to give rise to a large effect. However, for $x^+/x_P^+ \ll 1$ (close to the horizon), the shift in $\lambda\sigma^+$ is of order M/λ , an extremely large number. This implies that the shift is macroscopic in the sense that, for example, matter falling into the black hole some fixed time after the shock wave will end up at very different points on Σ , depending on whether we work in \mathcal{M} or $\bar{\mathcal{M}}$. Similarly, identical quantum states on \mathcal{I}_R^- should appear very different on Σ in the two cases, meaning that the matter state on Σ is strongly correlated with the fluctuations in geometry.

In this section, we will attempt to make the notion of different quantum states of matter on Σ more precise, allowing us to estimate the scale of entanglement between the matter and spacetime degrees of freedom. In order to do this, it is necessary to have a criterion to quantify the difference between two semiclassical matter states living in different spacetimes \mathcal{M} and $\bar{\mathcal{M}}$, that are identical on \mathcal{I}^- and are then evolved to Σ . The heuristic arguments above show that the expectation values of local operators can be very different for states in \mathcal{M} and $\bar{\mathcal{M}}$ that appear identical on \mathcal{I}^- where there is a fixed coordinate system through which to compare them. Rather than look at expectation values of operators, we construct an inner product

$$\langle \psi_1, \Sigma, \mathcal{M} | \psi_2, \Sigma, \bar{\mathcal{M}} \rangle$$

between Schrödinger picture matter states on the same Σ through which states on \mathcal{M} and $\bar{\mathcal{M}}$ can be compared. The inner product makes use of a decomposition in modes defined using the diffeomorphism invariant proper distance along Σ , through which the states can be compared. Details of this construction can be found in Appendix A.

An important feature of the inner product is that for a Planck scale fluctuation ΔM and for states $|\psi, \mathcal{M}\rangle$ and $|\psi, \bar{\mathcal{M}}\rangle$ that are identical on \mathcal{I}^- it can be checked that

$$\langle \psi, \Sigma, \mathcal{M} | \psi, \Sigma, \bar{\mathcal{M}} \rangle \approx 1 \quad (21)$$

on any generic surface Σ that does not have a large shift. This is a necessary condition for the consistency of quantum field theory on a mean curved background with a mass in the range $(M, M + \Delta M)$: If states on \mathcal{M} and $\bar{\mathcal{M}}$ are orthogonal on Σ , this is an indication that the approximate Hilbert space structure of the semiclassical approximation is becoming blurred due to an entanglement between the matter and gravity degrees of freedom. Using the inner product, we now show that matter states become approximately orthogonal on S-surfaces for extremely small fluctuations $\Delta M/\lambda \sim e^{-4\pi M/\lambda}$ in the mass of a black hole, dramatically violating condition (21).

In general the states that we wish to compare are most easily expressed as Heisenberg picture states on \mathcal{M} and $\bar{\mathcal{M}}$, and the prospect of converting these to Schrödinger picture states, and evolving them to Σ is rather daunting. As explained in Appendix A, there is a short cut to this procedure. For the states we are interested in (those that start as vacua on \mathcal{I}^-) the basic information needed for the calculation of the inner product is the relation induced by $\phi(s)$ on Σ between the tortoise coordinates on \mathcal{M} and $\bar{\mathcal{M}}$, namely $\sigma^+ = \sigma^+(\bar{\sigma}^+)$. This allows us to compute the inner product between the Schrödinger picture states by computing the usual Fock space inner product between two different Heisenberg picture states, defined with respect to the modes $e^{-i\omega\sigma^+}$ and $e^{-i\omega\bar{\sigma}^+}$. The latter inner product is given in terms of Bogoliubov coefficients. It should be stressed that this is just a short cut, and that the inner product depends crucially on the surface Σ , which is seen in the form of the function $\sigma^+ = \sigma^+(\bar{\sigma}^+)$.

We will study the overlap

$$\langle 0 \text{ in, } \Sigma, \mathcal{M} | 0 \text{ in, } \Sigma, \bar{\mathcal{M}} \rangle \quad (22)$$

where $|0 \text{ in, } \Sigma, \mathcal{M}\rangle$ is the matter Schrödinger picture state in spacetime \mathcal{M} on the hypersurface Σ which was in the natural left moving sector vacuum state on \mathcal{I}_R^- . We shall also use this quantity to estimate the size of $\Delta M = (\bar{M} - M)$ at which the states begin to differ appreciably. To evaluate the inner product (22), we first need to find the induced Bogoliubov transformation

$$v_\omega = \int_0^\infty d\omega' [\alpha_{\omega\omega'} \bar{v}_{\omega'} + \beta_{\omega\omega'} \bar{v}_{\omega'}^*] \quad (23)$$

between the in-modes

$$\begin{aligned} \bar{v}_\omega &= \frac{1}{\sqrt{2\omega}} e^{-i\omega\bar{\sigma}^+} \\ v_\omega &= \frac{1}{\sqrt{2\omega}} e^{-i\omega\sigma^+} , \end{aligned} \quad (24)$$

where σ^+ and $\bar{\sigma}^+$ are related by an induced relation

$$\sigma^+ = \sigma^+(\bar{\sigma}^+) . \quad (25)$$

Let us derive the relation (25) above, for the example of Section 3.3. As (20) shows us, the shift $\lambda\bar{x}^+ - \lambda x^+$ can become large and λx^+ above the shock wave maps to $\lambda\bar{x}^+$ further above³ the shock wave. As λx^+ comes closer to the shock wave and crosses to the other side, the image point $\lambda\bar{x}^+$ can still be located above the shock wave. Only when λx^+ is low enough under the shock wave, does $\lambda\bar{x}^+$ also cross the shock and go

³For simplicity, we will consider only the case $\Delta M < 0$ in this section. The conclusions will not depend on this assumption.

below it. Thus, the relation between the coordinates is split into three regions:

$$e^{-\phi} = \begin{cases} \alpha \lambda x^+ + \sqrt{\frac{M}{\lambda}} = \alpha \lambda \bar{x}^+ + \sqrt{\frac{M}{\lambda}} & (\lambda x^+ \geq 1) \\ \left(\alpha + \sqrt{\frac{M}{\lambda}} \right) \lambda x^+ = \alpha \lambda \bar{x}^+ + \sqrt{\frac{M}{\lambda}} & \left(1 \geq \lambda x^+ \geq \frac{\alpha + \sqrt{\frac{M}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}} \right) \\ \left(\alpha + \sqrt{\frac{M}{\lambda}} \right) \lambda x^+ = \left(\alpha + \sqrt{\frac{M}{\lambda}} \right) \lambda \bar{x}^+ & \left(\frac{\alpha + \sqrt{\frac{M}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}} \geq \lambda x^+ \geq 0 \right) \end{cases} . \quad (26)$$

Rewriting (26) using the asymptotic coordinates, we then get the relation (25):

$$\lambda \sigma^+ = \begin{cases} \ln \left[e^{\lambda \bar{\sigma}^+} - \frac{1}{\alpha} \left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{M}{\lambda}} \right) \right] & (\lambda \bar{\sigma}^+ \geq \lambda \bar{\sigma}_1^+) \\ \ln \left[\left(\frac{\alpha}{\alpha + \sqrt{\frac{M}{\lambda}}} \right) \left(e^{\lambda \bar{\sigma}^+} + \frac{1}{\alpha} \sqrt{\frac{M}{\lambda}} \right) \right] & (\lambda \bar{\sigma}_1^+ \geq \lambda \sigma^+ \geq 0) \\ \lambda \bar{\sigma}^+ + \ln \left[\frac{\alpha + \sqrt{\frac{M}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}} \right] & (0 \geq \lambda \bar{\sigma}^+) \end{cases} , \quad (27)$$

where

$$\lambda \bar{\sigma}_1^+ \equiv \ln \left[1 + \frac{1}{\alpha} \left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{M}{\lambda}} \right) \right] . \quad (28)$$

This coordinate transformation is illustrated in Fig. 5. As can be seen, in the first region (which corresponds to both points being above the shock) the transformation is logarithmic. On the other hand, in the third region when both points are below the shock, the transformation is exactly linear. The form of the transformation for the interpolating region when the other point is above and the other point below the shock should not be taken very seriously, since it depends on the assumptions made on the distribution of the infalling matter. For a shock wave it looks like a sharp jump, but if we smear the distribution to have a width of *e.g.* a Planck length, the jump gets smoothened and the transformation becomes closer to a linear one.

The Bogoliubov coefficients are now found to be

$$\begin{aligned} \alpha_{\omega\omega'} &= \frac{1}{2\pi} \sqrt{\frac{\omega'}{\omega}} I_{\omega\omega'}^+ \\ \beta_{\omega\omega'} &= \frac{1}{2\pi} \sqrt{\frac{\omega'}{\omega}} I_{\omega\omega'}^- , \end{aligned} \quad (29)$$

where $I_{\omega\omega'}^\pm$ are the integrals

$$I_{\omega\omega'}^\pm \equiv \int_{-\infty}^{\infty} d\bar{\sigma}^+ e^{-i\omega\sigma^+(\bar{\sigma}^+) \pm i\omega'\bar{\sigma}^+} . \quad (30)$$

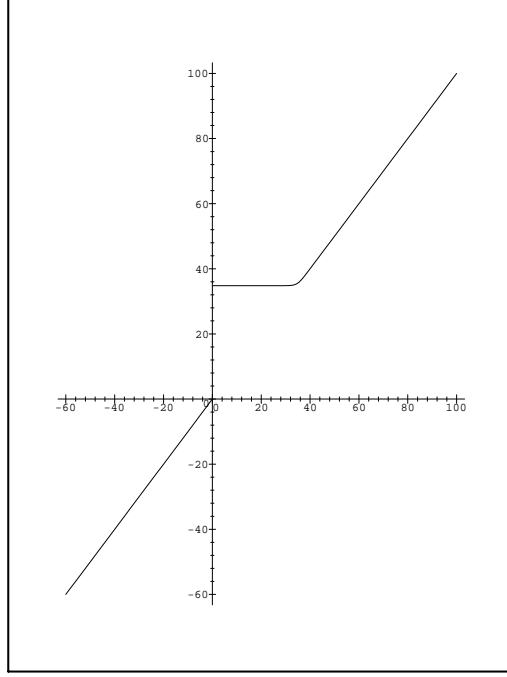


Figure 5: A graph of the function $\lambda\bar{\sigma}^+(\lambda\sigma^+)$. The vertical axis $=\lambda\bar{\sigma}^+$, the horizontal axis $=\lambda\sigma^+$. The part of the graph to the right (left) of the vertical axis corresponds to both points being above (below) the shock wave. (The interpolating part is in the region $\lambda\sigma^+ \in (-0.0013, 0)$ so the plot coincides with a segment of the vertical axis in the figure). The values $M/\lambda = 20$, $\Delta M/\lambda = -1/20$ and $r = 1/2$ were used in the plot.

Substituting the relations (27) above, we get

$$\begin{aligned}
I_{\omega\omega'}^{\pm} &= \int_{\bar{\sigma}_1^+}^{\infty} d\bar{\sigma}^+ \exp \left\{ -i\frac{\omega}{\lambda} \ln \left[e^{\lambda\bar{\sigma}^+} - \frac{1}{\alpha} \left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{M}{\lambda}} \right) \right] \pm i\omega'\bar{\sigma}^+ \right\} \\
&+ \int_0^{\bar{\sigma}_1^+} d\bar{\sigma}^+ \exp \left\{ -i\frac{\omega}{\lambda} \ln \left[\frac{\alpha}{\alpha + \sqrt{\frac{M}{\lambda}}} \left(e^{\lambda\bar{\sigma}^+} + \frac{1}{\alpha} \sqrt{\frac{M}{\lambda}} \right) \right] \pm i\omega'\bar{\sigma}^+ \right\} \\
&+ \int_{-\infty}^0 d\bar{\sigma}^+ \exp \left\{ -i\omega\bar{\sigma}^+ - i\frac{\omega}{\lambda} \ln \left[\frac{\alpha + \sqrt{\frac{M}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}} \right] \pm i\omega'\bar{\sigma}^+ \right\} \\
&\equiv I_{\omega\omega'}^{\pm}(1) + I_{\omega\omega'}^{\pm}(2) + I_{\omega\omega'}^{\pm}(3) \quad .
\end{aligned} \tag{31}$$

To identify the first integral, introduce first

$$\lambda\Delta \equiv 1 + \frac{1}{\alpha} \left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{M}{\lambda}} \right) = e^{\lambda\bar{\sigma}_1^+} . \tag{32}$$

and change the integration variable from $\lambda\bar{\sigma}^+$ to λy ,

$$\lambda\Delta e^{-\lambda y} \equiv e^{\lambda\bar{\sigma}^+} . \quad (33)$$

We now find that

$$I_{\omega\omega'}^{\pm}(1) = \lambda\Delta^{\pm i\omega'/\lambda} \int_{-\infty}^0 dy \exp \left\{ -i\frac{\omega}{\lambda} \ln [\lambda\Delta(e^{-\lambda y} - C)] \mp i\omega'y \right\} , \quad (34)$$

where

$$C \equiv 1 - \frac{1}{\lambda\Delta} . \quad (35)$$

Now we can recognize the integral to be the same as discussed in [14]. This integral can be identified as an incomplete β -function. However, it is also possible to make the standard approximation of replacing the integrand by its approximate value in the interval $\lambda y \in (-1, 0)$ [1, 14, 27]. Note that this interval corresponds to a region $\lambda\sigma^+ \in (0, \ln[(e-1)\lambda\Delta+1])$. The latter can be large: for $\lambda\Delta \sim e^{M/\lambda}$ it has size $\sim M/\lambda$. Indeed, comparing with (20) we notice that $\lambda\Delta - 1$ is equal to the shift $\lambda\bar{x}^+ - \lambda x^+$ above the shock, which could become exponential. So we can use

$$I_{\omega\omega'}^{\pm}(1) \approx \lambda\Delta^{\pm i\omega'/\lambda} \int_{-\infty}^0 dy \exp \left\{ -i\frac{\omega}{\lambda} \ln [-\lambda\Delta\lambda y] \mp i\omega'y \right\} . \quad (36)$$

As discussed in [27], this leads to the approximate relation

$$I_{\omega\omega'}^+(1) \approx -e^{\pi\omega/\lambda} \left(I_{\omega\omega'}^-(1) \right)^* \quad (37)$$

for the integrals.

The logarithm in (36) implies that $I_{\omega\omega'}^{\pm}(1)$ contributes significantly in the regime $\omega' - \omega \gg 1$. Since the coordinate transformation (27) was exactly linear in the third region and we argued that smearing of the incoming matter distribution smoothens the “interpolating part” in the second region, we can argue that $I_{\omega\omega'}^{\pm}(2)$ and $I_{\omega\omega'}^{\pm}(3)$ are negligible in the regime $\omega' - \omega \gg 1$. Therefore, in this limit $I_{\omega\omega'}^{\pm}(1)$ is the significant contribution, and a consequence of (37) is that the relationship of the Bogoliubov coefficients is (approximately) thermal,

$$\alpha_{\omega\omega'} \approx -e^{\pi\omega/\lambda} \beta_{\omega\omega'}^* , \quad (38)$$

with “temperature” $\lambda/2\pi$. Let us emphasize that the “temperature” itself is independent of the magnitude of the fluctuation ΔM . Rather, it is the *validity* of the thermal approximation that is affected: the larger the fluctuation ΔM is, the better approximation (38) is. Also, the region of $\lambda\sigma^+$ which corresponds to (38) becomes larger. Consequently, the inner product between $|0 \text{ in } \Sigma, \bar{\mathcal{M}}\rangle$ and $|0 \text{ in } \Sigma, \mathcal{M}\rangle$ can become appreciably smaller than 1. We refer to this as the states being “approximately orthogonal”, we will elaborate this below.

Let us now calculate the inner product (22). As was explained before, we have related this inner product to an inner product between two Heisenberg picture states

related by the above derived Bogoliubov transformation. For the latter inner product, we can use the general formula given in [30]. We then find (see Appendices):

$$|\langle 0 \text{ in, } \Sigma, \mathcal{M} | 0 \text{ in, } \Sigma, \bar{\mathcal{M}} \rangle|^2 = \left(\det(1 + \beta\beta^\dagger) \right)^{-\frac{1}{2}} . \quad (39)$$

We can now make an estimate of the scale of the fluctuations for the onset of the approximate orthogonality. As a rough criterion, let us say that as

$$|\langle 0 \text{ in, } \Sigma, \mathcal{M} | 0 \text{ in, } \Sigma, \bar{\mathcal{M}} \rangle|^2 < \frac{1}{\gamma} , \quad (40)$$

where $\gamma \sim e$, the states become approximately orthogonal. As is shown in Appendix B, the states become approximately orthogonal if

$$\left| \frac{\Delta M}{\lambda} \right| = \left| \frac{M}{\lambda} - \frac{\bar{M}}{\lambda} \right| > \gamma^{48} \sqrt{\frac{M}{\lambda}} \alpha , \quad (41)$$

where α is the (square root of the) slope. If the lines do not catch the Hawking radiation, $\alpha > 1$, then the fluctuations are not large enough to give arise to the approximate orthogonality and therefore (21) is satisfied. On the other hand, if the lines catch the fraction r of the Hawking radiation, $\alpha \approx e^{-4rM/\lambda}$ and the fluctuations can easily exceed the limit. (Recall that $M/\lambda \gg 1$, so α is the significant factor.) Note that the criterion (41) has been derived for the example hypersurfaces of section 3.3. However, the more general result for any S-surface of the types 1-3 of section 3.2 can be derived equally easily. In general the right hand side of (41) will depend on both the slope α^2 and the intercept⁴ δ . The physics of the result will remain the same as above: for the S-surfaces the approximate orthogonality begins as the fluctuations $\Delta M/\lambda$ satisfy $\ln(\lambda/\Delta M) \sim M/\lambda$.

One might ask what happens to the “in”-vacua at \mathcal{I}_L^- related to the rightmoving modes $e^{-i\omega y^-}, e^{-i\omega \bar{y}^-}$. We can similarly derive the induced coordinate transformations between the coordinates $\lambda \bar{y}^-$ and λy^- . This coordinate relation is virtually linear, and therefore the β Bogoliubov coefficients will be ≈ 0 and the vacua will have overlap ≈ 1 . Thus the effect is not manifest in the rightmoving sector.

5 Conclusions

Let us review what we have computed in this paper:

It is widely believed that the semiclassical approximation to gravity holds at the horizon of a black hole. We have computed a quantity that is natural in the consideration of the black hole problem, and that does *not* behave semiclassically at the horizon of the black hole. This quantity is the quantum state of matter on a hypersurface which

⁴Recall that the hypersurfaces of section 3.3 had $\delta = \alpha^2 + 2\alpha\sqrt{M/\lambda}$ so the rhs of (41) depends only on α .

also catches the outgoing Hawking radiation. The crucial ingredient of our approach was that when we try to get the semiclassical approximation from the full theory of quantum gravity, the natural quantity to compare between different semiclassical spacetimes is the same 1-geometry, not the hypersurface given by some coordinate relation on the semiclassical spacetimes. By contrast, in most calculations done with quantum gravity being a field theory on a background two dimensional spacetime, one computes n -point Greens functions, where the ‘points’ are given by chosen coordinate values in some coordinate system. For physics in most spacetimes, the answers would not differ significantly by either method, but in the presence of a black hole the difference is important.

We computed the quantum state on an entire spacelike hypersurface which goes up in time to capture the Hawking radiation, but then comes down steeply to intersect the infalling matter in the weak coupling region near the horizon. We found that quantum fluctuations in the background geometry prevent us from defining an unambiguous state on this S-surface. Matter states defined on an S-surface, evolved from a vacuum state at \mathcal{I}^- , are approximately orthogonal for fluctuations in mass of order e^{-M} or greater, a number much smaller than the size of fluctuations expected on general grounds.

One can expand the solution of the Wheeler–DeWitt equation in a different basis, such that for each term in this basis the total mass inside the hole is very sharply defined. If one ignores the Hawking radiation, then one finds that for such sharply defined mass the infalling matter has a large position uncertainty and cannot fall into the hole. Thus one may say that if one wants a good matter state on the S-slice, then the black hole cannot form. Any attempt to isolate a classical description for the metric while examining the quantum state for the matter will be impossible because the ‘gravity’ and matter modes are highly entangled. It is interesting that if we try to average over the ‘gravity states’ involved in the range $M \rightarrow M + \Delta M$, we generate entanglement entropy between ‘gravity’ and matter. This entropy is comparable to the entanglement entropy of Hawking pairs.

The computations of sections 3 and 4 show that the states on an S-surface differ appreciably in the region around the horizon. However, to calculate any local quantity close to the horizon, we could equally well have computed the state on spacelike hypersurfaces passing through the horizon without reaching up to \mathcal{I}^+ . On these surfaces we would find an unambiguous state of matter for black holes with masses differing on the Planck scale. This feature may signal that an effective theory of black hole evaporation might not be diffeomorphism invariant in the usual way. It also indicates that the breakdown in the semiclassical approximation is relevant only if we try to detect *both* the Hawking radiation and the infalling matter. Susskind has pointed to a possible complementarity between the description of matter outside the hole and the description inside. ’t Hooft and Schoutens *et. al.* have expressed this in terms of large commutators between operators localized at \mathcal{I}^+ and those localized close to the horizon. These notions of complementarity seem to be compatible with our results. It is interesting that we have arrived at them with minimal assumptions about the details of a quantum theory of gravity.

It should be mentioned that although every spacelike hypersurface that captures the Hawking radiation and the infalling matter near the horizon gives rise to the effect we have described, a slice that catches the Hawking radiation, enters the horizon high up, and catches the infalling matter deep inside the horizon can be seen to avoid the large shift. It seems that the quantum state of matter should be well defined on such a slice. The significance of this special case is not yet clear to us, although it is interesting that this type of slice appears to catch not only the Hawking pairs outside the horizon, but also their partners behind the horizon.

Our overall conclusion is that one must consider the entire solution of the quantum gravity problem near a black hole horizon, in particular one must take the solution to the Wheeler–DeWitt equation rather than its semiclassical projection. We believe that the arguments we have presented can be applied equally well to black holes in any number of dimensions.

Acknowledgements

We would like to thank R. Brooks, M. Crescimanno, J. Cohn, A. Dabholkar, S. Das, S. P. de Alwis, J. Goldstone, A. Guth, J. J. Halliwell, B. L. Hu, R. Jackiw, P. Kelly, K. Kuchař, G. Kunstatter, J. Samuel, A. Sen, C. Stephens, L. Susskind, S. Trivedi, H. Verlinde, A. Vilenkin, B. Zwiebach and especially J. G. Demers for discussions.

Appendix A

In this appendix we shall explain the construction of a natural inner product relating states of a quantum field defined on different spacetimes in which the same hypersurface Σ is embedded. Clearly the Schrödinger picture allows us to compute the value of a state on any hypersurface Σ in the form

$$\Psi[f(x), t_0]$$

where in the chosen coordinate system Σ is the surface $t = t_0$. There is of course a Hamiltonian operator which is coordinate dependent, and which in the chosen coordinate system specifies time evolution on constant t hypersurfaces:

$$H[f(x), \pi_f(x), x, t]\Psi = i\dot{\Psi}$$

We shall assume that the evolution of a state is independent of the coordinate system used, in the sense that a state on a Cauchy surface Σ_1 is taken to define a *unique* state on a later surface Σ_2 . This may not always be the case [28], but we will ignore such problems in our reasoning.

In quantum field theory on a fixed background, there is an inner product on the space of states on any hypersurface Σ . However, in order to use this inner product to compare states in different spacetimes, it is necessary to find a natural way of relating

two states defined on Σ without reference to coordinates. A natural way of doing this is to use the proper distance along Σ to define a mode decomposition, and to compare the states with respect to this decomposition.

Define

$$\begin{aligned} a(k) &= \int \frac{ds}{\sqrt{4\pi\omega_k}} e^{iks} \left(\omega_k f(x(s)) + \frac{\delta}{\delta f(x(s))} \right) \\ a^\dagger(k) &= \int \frac{ds}{\sqrt{4\pi\omega_k}} e^{-iks} \left(\omega_k f(x(s)) - \frac{\delta}{\delta f(x(s))} \right) \end{aligned} \quad (42)$$

so that $[a(k), a^\dagger(k')] = \delta(k - k')$. It is then straightforward to define the familiar inner product on the corresponding Fock space. An easy way to picture the Fock space in this Schrödinger picture language is to transform to a representation $\Psi[a^\dagger(k), t]$, in which $a(k)$ is represented as $\delta/\delta a^\dagger(k)$. The “vacuum” state, annihilated by all the $a(k)$, is just the functional $\Psi = 1$, and excited states arise from multiplication by $a^\dagger(k)$. The inner product is

$$(\Psi_1, \Psi_2) = \int \frac{\prod_k da(k) da^\dagger(k)}{2\pi i} \left(\Psi_1[a^\dagger(k), t] \right)^* \Psi_2[a^\dagger(k), t] \exp \left[- \int dk a(k) a^\dagger(k) \right] \quad (43)$$

where $(a^\dagger(k))^* = a(k)$ [29].

We could carry out the same construction for any coordinate x on Σ , and the inner products would necessarily agree. However, the operators $a(k)$ and $a^\dagger(k)$ constructed using proper distance are special, in that we shall say that two states $\Psi_{\mathcal{M}}[f(x), t_1]$ and $\Psi_{\bar{\mathcal{M}}}[f(x), t_2]$ defined on different spacetimes \mathcal{M} and $\bar{\mathcal{M}}$ but on a common hypersurface located at $t_{\mathcal{M}} = t_1$ or $t_{\bar{\mathcal{M}}} = t_2$, are the same if they are the same Fock states with respect to this decomposition. If they are not identical, their overlap is given by the Fock space inner product (43), and this is what is meant in Section 4 by

$$\langle \psi_1, \Sigma, \mathcal{M} | \psi_2, \Sigma, \bar{\mathcal{M}} \rangle.$$

Having defined an inner product between two Schrödinger picture states on different spacetimes, we want to extend it to Heisenberg picture states on \mathcal{M} and $\bar{\mathcal{M}}$, since these are the kind of states we are used to working with in quantum field theory in curved spacetime. The inner product we have just defined can be used to relate Heisenberg picture states by transforming each of the states to the Schrödinger picture, evolving them to the common hypersurface Σ , and computing the overlap there. It is useful to have a short-cut to this computation. In order to achieve this, we first relate a Schrödinger picture state $\Psi[a^\dagger(k), t_0]$ to a Heisenberg picture state $\Psi[a^\dagger(k)]$, where now the $a(k)$ are associated with mode functions on \mathcal{M} not Σ : First choose coordinates (x, t) on the spacetime such that the metric is conformally flat, that Σ is a constant time slice, $t = t_0$, and that the conformal factor is unity on Σ . Using these coordinates, we can compute the Hamiltonian, which by virtue of two dimensional conformal invariance is free

$$H = \int dx (\pi_f^2 + (f')^2).$$

We pick a mode basis defined by

$$\begin{aligned} a(k) &= \int \frac{dx}{\sqrt{4\pi\omega_k}} e^{ikx} \left(\omega_k f(x) + \frac{\delta}{\delta f(x)} \right) \\ a^\dagger(k) &= \int \frac{dx}{\sqrt{4\pi\omega_k}} e^{-ikx} \left(\omega_k f(x) - \frac{\delta}{\delta f(x)} \right) \end{aligned}$$

so that on Σ this is precisely the proper distance mode decomposition. In terms of these modes, the Hamiltonian is simply given by

$$H = \int dk \omega_k a^\dagger(k) a(k)$$

so that transforming the operators $a(k)$ and $a^\dagger(k)$ to the Heisenberg picture simply gives

$$a(k, t) = e^{i\omega_k(t-t_0)} a(k), \quad a^\dagger(k, t) = e^{-i\omega_k(t-t_0)} a^\dagger(k)$$

and

$$f(x, t) = \int \frac{dk}{\sqrt{4\pi\omega_k}} \left(a(k) e^{-ik \cdot x} + a^\dagger(k) e^{ik \cdot x} \right).$$

Correspondingly the Schrödinger picture state $\Psi[a^\dagger(k), t]$ is identical in form to the Heisenberg picture state: $\Psi[a^\dagger(k)] \equiv \Psi[a^\dagger(k), t]|_{t=t_0}$. We may repeat this procedure on another spacetime $\bar{\mathcal{M}}$, again defining modes on $\bar{\mathcal{M}}$ so that $\Psi_{\bar{\mathcal{M}}}[a^\dagger(k)]$ is identical to the Schrödinger picture state on $\Sigma_{\bar{\mathcal{M}}}$. Then, the inner product (43) serves as an inner product for Heisenberg picture states $\Psi_{\mathcal{M}}[a^\dagger(k)]$ and $\Psi_{\bar{\mathcal{M}}}[a^\dagger(k)]$ living on spacetimes \mathcal{M} and $\bar{\mathcal{M}}$.

Now in order to compare a Heisenberg picture state $\Psi_{\mathcal{M}}[a^\dagger(k)]$ to any other Heisenberg picture state on \mathcal{M} , we may make use of standard Bogoliubov coefficient techniques. Consider another state defined in the Heisenberg picture in terms of mode-coefficients related to modes $v_p(x, t)$ on \mathcal{M} . Let us suppose that we associate operators $b(p)$ and $b^\dagger(p)$ with the modes $v_p(x, t)$. Then

$$\begin{aligned} b(p) &= \sum_k \left[\alpha_{kp} a(k) + \beta_{kp}^* a^\dagger(k) \right] \\ b^\dagger(p) &= \sum_k \left[\beta_{kp} a(k) + \alpha_{kp}^* a^\dagger(k) \right] \end{aligned} \tag{44}$$

where

$$\alpha_{kp} = -i \int dx f_k(x, t) \partial_t v_p^*(x, t), \quad \beta_{kp} = i \int dx f_k(x, t) \partial_t v_p(x, t).$$

Here $f_k(x, t) = e^{-ik \cdot x} / \sqrt{4\pi\omega_k}$ are the modes defining the $a(k)$.

We may perform a similar calculation on a neighbouring spacetime $\bar{\mathcal{M}}$, also containing Σ , to relate a set of modes $\bar{v}_q(\bar{x}, \bar{t})$ to the modes $f_k(\bar{x}, \bar{t})$ and similarly to relate the operators $a(k)$ and $a^\dagger(k)$ to the $\bar{b}(p)$ and $\bar{b}^\dagger(p)$ as

$$\begin{aligned} \bar{b}(p) &= \sum_k \left[\bar{\alpha}_{kp} a(k) + \bar{\beta}_{kp}^* a^\dagger(k) \right] \\ \bar{b}^\dagger(p) &= \sum_k \left[\bar{\beta}_{kp} a(k) + \bar{\alpha}_{kp}^* a^\dagger(k) \right] \end{aligned} \tag{45}$$

Now we can use the inner product (43) to relate two states $\Psi_{\mathcal{M}}[b^\dagger(p)]$ and $\Psi_{\bar{\mathcal{M}}}[\bar{b}^\dagger(p')]$ directly.

More simply, it follows from (44) and (45) that the b 's and \bar{b} 's are related by

$$\begin{aligned} b(p') &= \sum_p \left[(\bar{v}_p, v_{p'}) \bar{b}(p) + (\bar{v}_p^*, v_{p'}) \bar{b}^\dagger(p) \right] \\ b^\dagger(p') &= \sum_p \left[-(\bar{v}_p^*, v_{p'}) \bar{b}^\dagger(p) - (\bar{v}_p, v_{p'}) \bar{b}(p) \right] \end{aligned} \quad (46)$$

where

$$(\bar{v}_p, v_{p'}) = -i \int_{\Sigma} dx \bar{v}_p(x, t) \partial_t v_{p'}^*(x, t) \quad (47)$$

so that the inner product between states on \mathcal{M} and $\bar{\mathcal{M}}$ may be computed using the standard inner product for states $\Psi[b^\dagger(p)]$ without going through the $a(k)$.

In the examples that we consider, the Bogoliubov coefficients in (46) need not be evaluated on Σ as in (47). Suppose for example that we have left moving mode bases $v_p(\sigma^+)$ and $\bar{v}_p(\bar{\sigma}^+)$ defined in terms of tortoise coordinates on \mathcal{M} and $\bar{\mathcal{M}}$ respectively. Then both v_p and \bar{v}_p are functions of x^+ only. We can simply change variables in (47) from x to σ (the t differentiation becomes an x differentiation which absorbs the factor $dx/d\sigma$), yielding

$$(\bar{v}_p, v_{p'}) = -i \int d\sigma \bar{v}_p(\bar{\sigma}^+(\sigma^+)) \partial_{\sigma^0} v_{p'}^*(\sigma^+) \quad (48)$$

where $\bar{\sigma}^+$ is given as a function σ^+ through the relations derived by equating points on Σ according to the values of ϕ and $d\phi/ds$, as in Section 3. The integral (48) looks exactly like the familiar integral for Bogoliubov coefficients, even though it involves mode functions on different manifolds. (48) may be evaluated on any Cauchy surface in \mathcal{M} (or $\bar{\mathcal{M}}$) since both the mode functions solve the Klein-Gordon equation on \mathcal{M} ($\bar{\mathcal{M}}$).

Appendix B

We present some details of the calculation of the overlap of the two states on Σ . We now know the Bogoliubov transformation between the modes v_ω and \bar{v}_ω in the text. Subsequently, the overlap of the two Schrödinger picture states can be found to be

$$\langle 0 \text{ in } \Sigma, \mathcal{M} | 0 \text{ in } \Sigma, \bar{\mathcal{M}} \rangle = (\det(\alpha))^{-\frac{1}{2}}, \quad (49)$$

where α is the matrix $(\alpha_{\omega\omega'})$ of Bogoliubov coefficients. The right hand side is the general formula for the overlap of two vacuum states related to modes connected by a Bogoliubov transformation [30]. However, it is more convenient to consider not the overlap but the probability amplitude

$$|\langle 0 \text{ in } \Sigma, \mathcal{M} | 0 \text{ in } \Sigma, \bar{\mathcal{M}} \rangle|^2 = (\det(\alpha\alpha^\dagger))^{-\frac{1}{2}}, \quad (50)$$

where the components of the matrix $\alpha\alpha^\dagger$ are

$$(\alpha\alpha^\dagger)_{\omega\omega'} = \int_0^\infty d\omega'' \alpha_{\omega\omega''} \alpha_{\omega'\omega''}^* . \quad (51)$$

The evaluation of the determinant of the matrix $\alpha\alpha^\dagger$ becomes easier if we move into a wavepacket basis. Instead of the modes v_ω we use

$$v_{jn} \equiv a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega e^{2\pi i \omega n/a} v_\omega . \quad (52)$$

These wavepackets are centered at $\sigma^+ = 2\pi n/a$, where $n = \dots, -1, 0, 1, \dots$, they have spatial width $\sim a^{-1}$ and a frequency $\omega_j \approx ja$, where $j = 0, 1, \dots$. For more discussion, see [1, 27, 14]. In the new basis, the Bogoliubov coefficients become

$$\begin{aligned} \alpha_{jn\omega'} &= a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega e^{2\pi i \omega n/a} \alpha_{\omega\omega'} \\ \beta_{jn\omega'} &= a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega e^{2\pi i \omega n/a} \beta_{\omega\omega'} \end{aligned} \quad (53)$$

with the normalization

$$\int_0^\infty d\omega'' [\alpha_{jn\omega''} \alpha_{j'n'\omega''}^* - \beta_{jn\omega''} \beta_{j'n'\omega''}^*] = \delta_{jj'} \delta_{nn'} . \quad (54)$$

The thermal relation (38) becomes

$$\beta_{jn\omega'}^* \approx -e^{-\pi\omega_j/\lambda} \alpha_{jn\omega'} . \quad (55)$$

Recall that the validity of the thermal approximation corresponded to the region $\lambda\sigma^+ \in (0, \ln[1 + (e-1)\lambda\Delta])$ where $\lambda\Delta - 1$ was the shift $\lambda\bar{x}^+ - \lambda x^+$. Let us denote the size of this region as λL . Since the separation of the wavepackets is $\Delta(\lambda\sigma^+) = 2\pi\lambda/a$, we can say that

$$n_{max} = \frac{\lambda L}{\Delta(\lambda\sigma^+)} = \frac{\ln[1 + (e-1)\lambda\Delta]}{2\pi\lambda/a} \quad (56)$$

packets are centered in this region.

Combining (54) and (55), we now see that

$$(\alpha\alpha^\dagger)_{jnj'n'} \approx \frac{\delta_{jj'} \delta_{nn'}}{1 - e^{-2\pi\omega_j/\lambda}} \quad (57)$$

for n, n' “inside” λL . For the other values of n, n' (at least one of them being “outside”),

$$(\alpha\alpha^\dagger)_{jnj'n'} \approx \delta_{jj'} \delta_{nn'} . \quad (58)$$

We are now ready to calculate the overlap (50). We get

$$\begin{aligned} \ln [|\langle 0 \text{ in, } \Sigma, \mathcal{M} | 0 \text{ in, } \Sigma, \bar{\mathcal{M}} \rangle|^2] &\approx -\frac{1}{2} \left\{ \sum_n^{(\text{inside})} \sum_j \ln \left[\frac{1}{1 - e^{-2\pi\omega_j/\lambda}} \right] \right. \\ &\quad \left. + \ln \left[\prod_n^{(\text{outside})} \prod_j 1 \right] \right\} \\ &= \frac{1}{2} n_{max} \sum_j \ln [1 - e^{-2\pi ja/\lambda}] . \end{aligned} \quad (59)$$

In order to estimate the last term, we convert the sum to an integral⁵:

$$\begin{aligned} \sum_j \ln [1 - e^{-2\pi ja/\lambda}] &\rightarrow \int_0^\infty dj \ln [1 - e^{-2\pi ja/\lambda}] \\ &= \frac{\lambda}{2\pi a} \left(\frac{-\pi^2}{6} \right) . \end{aligned} \quad (60)$$

Now, combining (59) and (60), we finally get a useful formula for the overlap:

$$\begin{aligned} |\langle 0 \text{ in, } \Sigma, \mathcal{M} | 0 \text{ in, } \Sigma, \bar{\mathcal{M}} \rangle|^2 &\approx \exp \left\{ -\frac{1}{2} \frac{\lambda L}{2\pi\lambda/a} \frac{\lambda}{2\pi a} \frac{\pi^2}{6} \right\} \\ &= e^{-\frac{1}{48}\lambda L} . \end{aligned} \quad (61)$$

Now we can estimate when the overlap is $< \gamma^{-1}$ where γ is a number $\sim e$. The overlap becomes equal to γ^{-1} as

$$\begin{aligned} 48 \ln \gamma &= \lambda L = \ln [1 + (e - 1)\lambda\Delta] \\ &\approx \ln \left[1 + (e - 1) \left(1 - \frac{\Delta M}{2\alpha\lambda} \sqrt{\frac{\lambda}{M}} \right) \right] , \end{aligned} \quad (62)$$

where we used (20) in the last step. Solving for ΔM , we get

$$\frac{\Delta M}{\lambda} \approx -\frac{2}{e - 1} \gamma^{48} \sqrt{\frac{M}{\lambda}} \alpha . \quad (63)$$

If $|\Delta M/\lambda|$ is bigger, the states are approximately orthogonal. Notice that since we used (20) in the end, (63) is a special result for the hypersurfaces of section 3.3. However, it is straightforward to generalize (63) to any S-surface of section 3.2 by using the relevant shifts as $\lambda\Delta - 1$ and proceeding as above. In general the right hand side of (63) will then depend on both α and the intercept δ .

⁵Notice that one might like to exclude frequencies corresponding to wavelengths much larger than the thermal region λL and impose an infra-red cut off at $j_{min}a \sim 1/L$. It turns out that for $\infty > \lambda L > 2\pi$ ($0 < j_{min}a < \lambda/2\pi$) the effect of imposing this cut off is negligible. Therefore we can just as well take the integral over the full range.

References

- [1] S. W. Hawking, *Comm. Math. Phys.* **43** (1975) 199.
- [2] S. W. Hawking, *Phys. Rev.* **D14** (1976) 2460.
- [3] C. G. Callan, S. B. Giddings, J. A. Harvey and A. Strominger, *Phys. Rev. D* **45** (1992) R1005. For reviews see J. A. Harvey and A. Strominger, *Quantum Aspects of Black Holes*, in the proceedings of the 1992 TASI Summer School in Boulder, Colorado (World Scientific, 1993), and S. B. Giddings, *Toy Models for Black Hole Evaporation*, in the proceedings of the International Workshop of Theoretical Physics, 6th Session, June 1992, Erice, Italy, ed. V. Sanchez (World Scientific, 1993).
- [4] G. 't Hooft, *Nucl. Phys.* **B335** (1990) 138, and references therein.
- [5] G. 't Hooft, *Nucl. Phys.* **B256** (1985) 727.
- [6] C. R. Stephens, G. 't Hooft and B. F. Whiting, *Class. Qu. Grav.* **11** (1994) 621.
- [7] T. Jacobson, *Phys. Rev.* **D48** (1993) 728.
- [8] D. N. Page, *Phys. Rev. Lett.* **44** (1980) 301.
- [9] L. Susskind, L. Thorlacius and J. Uglum, *Phys. Rev.* **D48** (1993) 3743.
- [10] L. Susskind, *Phys. Rev.* **D49** (1994) 6606.
- [11] E. Verlinde and H. Verlinde, *A Unitary S-matrix for 2D Black Hole Formation and Evaporation*, Princeton Preprint, PUPT-1380, IASSNS-HEP-93/8, hep-th/9302022 (1993), K. Schoutens, E. Verlinde, and H. Verlinde, *Phys. Rev.* **D48** (1993) 2670.
- [12] S. R. Das and S. Mukherji, *Phys. Rev.* **D50** (1994) 930; A. Strominger and L. Thorlacius, *Conformally Invariant Boundary Conditions for Dilaton Gravity*, Univ. of California at Santa Barbara report No. NSF-ITP-94-34, (hep-th/9405084).
- [13] A. Bilal and C. Callan, *Nucl. Phys.* **B394** (1993) 73; K. Nakamura, S. Konno, Y. Oshiro and A. Tomimatsu, *Prog. Th. Phys.* **90** (1993) 861; Y. Kazama and Y. Satoh, *Extraction of Black Hole Geometry in Exactly Quantized Two-Dimensional Dilaton Gravity*, Tokyo U. Report UT-Komoba-93-13 (hep-th/9310155); *Space-Time Geometry in Exactly Solvable Quantum Dilaton Gravity*, Tokyo U. Report UT-Komoba-93-13 (hep-th/9310155); *Phys. Rev.* **D48** (1993) 1687; S. P. de Alwis, *Phys. Lett.* **B317** (1993) 46; S. P. de Alwis and D. A. MacIntire, *The Problem of Time and Quantum Black Holes*, Colorado U. Report COLO-HEP-333 (gr-qc/9403052); C. Kiefer, R. Muller and T. P. Singh, *Quantum Gravity and Non-Unitarity in Black Hole Evaporation*, Zurich U. Report

- ZU-TH-25-93 (gr-qc/9308024); T. Hori, *Prog. Th. Phys.* **90** (1993) 743; T. Hori and M. Kamata, *Wave-Functional of Quantum Black Holes in Two-Dimensions*, Teikyo U. Report TEP-10 (hep-th/9302142).
- [14] E. Keski-Vakkuri and S. D. Mathur, *Phys. Rev.* **D50** (1994) 917.
 - [15] T. M. Fiola, J. Preskill, A. Strominger and S. P. Trivedi, *Black Hole Thermodynamics and Information Loss in Two Dimensions*, Caltech report n.o. CALT-68-1918, (hep-th/9403137); C. Holzhey, F. Larsen and F. Wilczek, *Geometric and Renormalized Entropy in Conformal Field Theory*, Princeton Report No. PUPT-1454, IASSNS-HEP 93/88 (hep-th/9403108).
 - [16] W. H. Zurek, *Phys. Rev. Lett.* **49** (1982) 1683.
 - [17] V. Lapchinsky and V. Rubakov, *Acta Phys. Pol.* **B10** (1979) 1041; T. Banks, *Nucl. Phys.* **B249** (1985) 332; J. Hartle, *Progress in Quantum Cosmology*, in *Proceedings, 12th Conference on General Relativity and Gravitation, Boulder, 1989*, edited by N. Ashby, D. F. Bartlett and W. Wyss (CUP, Cambridge, 1990); J. J. Halliwell, *Phys. Rev.* **D43** (1991) 2590; C. Kiefer, *The Semiclassical Approximation to Quantum Gravity*, Freiburg University Report No. THEP-93/27, to appear in *Canonical Gravity - from Classical to Quantum*, edited by J. Ehlers and H. Friedrich (Springer, Berlin 1994) (gr-qc/9312015).
 - [18] S. D. Mathur, *Black Hole Entropy and the Semiclassical Approximation*, MIT report No. CTP-2304 (hep-th/9404135) (Invited Talk given at the International Colloquium on Modern Quantum Field Theory II at TIFR (Bombay), January 1994).
 - [19] E. Calzetta and F. D. Mazzitelli, *Phys. Rev.* **D42** (1990) 4066.
 - [20] J. P. Paz and S. Sinha, *Phys. Rev.* **D44** (1991) 1038.
 - [21] P. A. M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School Monograph 2, Yeshiva University, NY, 1964; R. Arnowitt, S. Deser, C. Misner, in *Gravitation*, ed. L. Witten, Wiley, NY (1962) p. 227; B. DeWitt, *Phys. Rev.* **160** (1967) 1113; C. Misner, K. Thorne, J. Wheeler, *Gravitation*, W. H. Freeman, San Francisco (1971); A. Hanson, T. Regge, C. Teitelboim, *Constrained Hamiltonian Systems*, Lincei Academy report No. 2 (1976); C. J. Isham, Lectures given at 1983 Les Houches School on Relativity Groups and Topology, June 27 - Aug. 4 1983; *Proc. Roy. Soc.* **A368** (1979) 33; M. Pilati, *Phys. Rev.* **D26** (1982) 2645.
 - [22] See for example S. B. Giddings, *Phys. Rev.* **D49** (1994) 4078.
 - [23] J. G. Russo, L. Susskind and L. Thorlacius, *Phys. Rev.* **D46** (1992) 3444; *Phys. Rev.* **D47** (1993) 533.

- [24] J. Preskill, *Do black holes destroy information?*, CalTech Report No. CALT-68-1819 (hep-th/9209058);
- [25] S. Giddings, *Black holes and quantum predictability*, Santa Barbara Report No. UCSBTH-93-16 (hep-th/9306041).
- [26] T. Banks and M. O’Loughlin, *Nucl. Phys.* **B362** (1991) 649; D. Louis-Martinez, J. Gegenberg and G. Kunstatter, *Phys. Lett.* **B321** (1994) 193; K. V. Kuchař *Geometrodynamics of Schwarzschild Black Holes*, Utah Preprint UU-REL-94/3/1, (gr-qc/9403003); H. A. Kastrup and T. Thiemann, *Spherically Symmetric Gravity as a Completely Integrable System*, Preprint PITHA 93-35, (gr-qc/9401032).
- [27] S. B. Giddings and W. M. Nelson, *Phys. Rev.* **D46** (1992) 2486.
- [28] K. V. Kuchař, *Phys. Rev.* **D39** (1989) 1579.
- [29] L. D. Faddeev and A. A. Slavnov, *Gauge Fields: An Introduction to Quantum Theory*, Frontiers in Physics # 83, Addison-Wesley, Redwood City (1991).
- [30] B. S. DeWitt, *Phys. Rep.* **19C** (1975) 297.